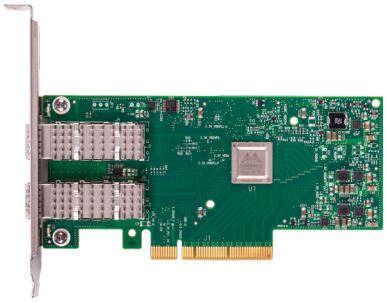
ADAPTER CARD
PRODUCT BRIEF



ConnectX®-4 Lx EN Card

50Gb/s Ethernet Adapter Card

ConnectX-4 Lx †

1/10/25/40/50 Gigabit Ethernet Adapter Cards supporting RDMA, Overlay Networks Encapsulation/Decapsulation and more

ConnectX-4 Lx EN Network Controller with 1/10/25/40/50Gb/s Ethernet connectivity addresses virtualized infrastructure challenges, delivering best-in-class and highest performance to various demanding markets and applications. Providing true hardware-based I/O isolation with unmatched scalability and efficiency, achieving the most cost-effective and flexible solution for Web 2.0, Cloud, data analytics, database, and storage platforms.

With the exponential increase in usage of data and the creation of new applications, the demand for the highest throughput, lowest latency, virtualization and sophisticated data acceleration engines continues to rise. ConnectX-4 Lx EN enables data centers to leverage the world's leading interconnect adapter for increasing their operational efficiency, improving servers' utilization, maximizing applications productivity, while reducing total cost of ownership (TCO).

ConnectX-4 Lx EN provides an unmatched combination of 1, 10, 25, 40, and 50GbE bandwidth, sub-microsecond latency and a 75 million packets per second message rate. It includes native hardware support for RDMA over Converged Ethernet, Ethernet stateless offload engines, Overlay Networks,and GPUDirect® Technology.

HIGH SPEED ETHERNET ADAPTER

ConnectX-4 Lx offers the best cost effective Ethernet adapter solution for 1, 10, 25, 40 and 50Gb/s Ethernet speeds, enabling seamless networking, clustering, or storage. The adapter reduces application runtime, and offers the flexibility and scalability to make infrastructure run as efficiently and productively as possible.

I/O VIRTUALIZATION

ConnectX-4 Lx EN SR-IOV technology provides dedicated adapter resources and guaranteed isolation and protection for virtual machines (VMs) within the server. I/O virtualization with ConnectX-4 Lx EN gives data center administrators better server utilization while reducing cost, power, and cable complexity, allowing more Virtual Machines and more tenants on the same hardware.

OVERLAY NETWORKS

In order to better scale their networks, data center operators often create overlay networks that carry traffic from individual virtual machines over logical tunnels in encapsulated formats such as NVGRE and VXLAN. While this solves network scalability issues, it hides the TCP packet from the hardware offloading engines, placing higher loads on the host CPU. ConnectX-4 Lx EN effectively addresses this by providing advanced NVGRE, VXLAN and GENEVE hardware offloading engines that encapsulate and

HIGHLIGHTS

NEW FEATURES

- 1/10/25/40/50Gb/s speeds
- Single and dual-port options available
- Erasure Coding offload
- Virtualization
- Accelerated Switching and Packet Processing (ASAP²)
- Low latency RDMA over Converged Ethernet (RoCE)
- CPU offloading of transport operations
- Application offloading
- Mellanox PeerDirect™ communication acceleration
- Hardware offloads for NVGRE, VXLAN and GENEVE encapsulated traffic
- End-to-end QoS and congestion control
- Hardware-based I/O virtualization
- RoHS compliant
- ODCC compatible

BENEFITS

- Highest performing boards for applications requiring high bandwidth, low latency and high message rate
- Industry leading throughput and latency for Web 2.0, Cloud and Big Data applications
- Smart interconnect for x86, Power, ARM, and GPU-based compute and storage platforms
- Cutting-edge performance in virtualized overlay networks
- Efficient I/O consolidation, lowering data center costs and complexity
- Virtualization acceleration
- Power efficiency

† For illustration only. Actual products may vary.

de-capsulate the overlay protocol headers, enabling the traditional offloads to be performed on the encapsulated traffic for these and other tunneling protocols (GENEVE, MPLS, QinQ, and so on). With ConnectX-4 Lx EN, data center operators can achieve native performance in the new network architecture.

ASAP^{2} ™

Mellanox ConnectX-4 Lx EN offers Accelerated Switching And Packet Processing (ASAP^{2}) technology to perform offload activities in the hypervisor, including data path, packet parsing, VxLAN and NVGRE encapsulation/decapsulation, and more.

ASAP^{2} allows offloading by handling the data plane in the NIC hardware using SR-IOV, while maintaining the control plane used in today's software-based solutions unmodified. As a result, there is significantly higher performance without the associated CPU load. ASAP^{2} has two formats: ASAP^{2} Flex™ and ASAP^{2} Direct™.

One example of a virtual switch that ASAP^{2} can offload is OpenVSwitch (OVS).

RDMA OVER CONVERGED ETHERNET (RoCE)

ConnectX-4 Lx EN supports RoCE specifications delivering low-latency and high- performance over Ethernet networks. Leveraging data center bridging (DCB) capabilities as well as ConnectX-4 Lx EN advanced congestion control hardware mechanisms, RoCE provides efficient low-latency RDMA services over Layer 2 and Layer 3 networks.

MELLANOX PEERDIRECT™

PeerDirect™ communication provides high efficiency RDMA access by eliminating unnecessary internal data copies between components on the PCIe bus (for example, from GPU to CPU), and therefore significantly reduces application run time. ConnectX-4 Lx EN advanced acceleration technology enables higher cluster efficiency and scalability to tens of thousands of nodes.

STORAGE ACCELERATION

Storage applications will see improved performance with the higher bandwidth ConnectX-4 Lx EN delivers. Moreover, standard block and file access protocols can leverage RoCE for high-performance storage access. A consolidated compute and storage network achieves significant cost-performance advantages over multi-fabric networks.

DISTRIBUTED RAID

ConnectX-4 Lx EN delivers advanced Erasure Coding offloading capability, enabling distributed Redundant Array of Inexpensive Disks (RAID), a data storage technology that combines multiple disk drive components into a logical unit for the purposes of data redundancy and performance improvement. ConnectX-4 Lx EN's Reed-Solomon capability introduces redundant block calculations, which, together with RDMA, achieves high performance and reliable storage access.

HOST MANAGEMENT

Mellanox host management and control capabilities include NC-SI over MCTP over SMBus, and MCTP over PCIe - Baseboard Management Controller (BMC) interface, as well as PLDM for Monitor and Control DSP0248 and PLDM for Firmware Update DSP0267.

SOFTWARE SUPPORT

All Mellanox adapter cards are supported by Windows, Linux distributions, VMware, FreeBSD, and Citrix XENServer. ConnectX-4 Lx EN supports various management interfaces and has a rich set of tools for configuration and management across operating systems.

COMPATIBILITY

PCI Express Interface

- PCIe Gen 3.0 compliant, 1.1 and 2.0 compatible
- 2.5, 5.0, or 8.0GT/s link rate x8
- Auto-negotiates to x8, x4, x2, or x1
- Support for MSI/MSI-X mechanisms

Operating Systems/Distributions*

- RHEL/CentOS
- Windows
- FreeBSD
- VMware
- OpenFabrics Enterprise Distribution (OFED)
- OpenFabrics Windows Distribution (WinOF-2)

Connectivity

- Interoperable with 1/10/25/40/50Gb Ethernet switches
- Passive copper cable with ESD protection
- Powered connectors for optical and active cable support

FEATURES

Ethernet
- 50GbE / 40GbE / 25GbE / 10GbE / 1GbE
- IEEE 802.3bj, 802.3bm 100 Gigabit Ethernet
- 25G Ethernet Consortium 25, 50 Gigabit Ethernet
- IEEE 802.3ba 40 Gigabit Ethernet
- IEEE 802.3ae 10 Gigabit Ethernet
- IEEE 802.3az Energy Efficient Ethernet
- IEEE 802.3ap based auto-negotiation and KR startup
- Proprietary Ethernet protocols (20/40GBASE-R2, 50GBASE-R4)
- EEE 802.3ad, 802.1AX Link Aggregation
- IEEE 802.1Q, 802.1P VLAN tags and priority
- IEEE 802.1Qau (QCN) – Congestion Notification
- EEE 802.1Qaz (ETS)
- IEEE 802.1Qbb (PFC)
- IEEE 802.1Qbg
- IEEE 1588v2
- Jumbo frame support (9.6KB)

Enhanced Features
- Hardware-based reliable transport
- Collective operations offloads
- Vector collective operations offloads
- PeerDirect RDMA (aka GPUDirect® communication acceleration
- 64/66 encoding
- Extended Reliable Connected transport (XRC)
- Dynamically Connected transport (DCT)
- Enhanced Atomic operations
- Advanced memory mapping support, allowing user mode registration and remapping of memory (UMR)
- On demand paging (ODP) – registration free RDMA memory access

CPU Offloads
- RDMA over Converged Ethernet (RoCE)
- TCP/UDP/IP stateless offload
- LSO, LRO, checksum offload
- RSS (can be done on encapsulated packet), TSS, HDS, VLAN insertion/ stripping, Receive flow steering
- Intelligent interrupt coalescence

Storage Offloads
- RAID offload - erasure coding (Reed-Solomon) offload

Overlay Networks
- Stateless offloads for overlay networks and tunneling protocols
- Hardware offload of encapsulation and decapsulation of NVGRE and VXLAN overlay networks

Hardware-Based I/O Virtualization
- Single Root IOV
- Multi-function per port
- Address translation and protection
- Multiple queues per virtual machine
- Enhanced QoS for vNICs
- VMware NetQueue support

Virtualization
- SR-IOV: Up to 256 Virtual Functions
- SR-IOV: Up to 16 Physical Functions per port
- Virtualization hierarchies (e.g. NPAR)
  - Virtualizing Physical Functions on a physical port
  - SR-IOV on every Physical Function
- 1K ingress and egress QoS levels
- Guaranteed QoS for VMs

Protocol Support
- OpenMPI, IBM PE, OSU MPI (MVAPICH/2), Intel MPI
- Platform MPI, UPC, Open SHMEM
- TCP/UDP, MPLS, VxLAN, NVGRE, GENEVE
- iSER, NFS RDMA, SMB Direct
- uDAPL

Management and Control Interfaces
- NC-SI over MCTP over SMBus and NC-SI over MCTP over PCIe - Baseboard Management Controller interface
- SDN management interface for managing the eSwitch
- I^2C interface for device control and configuration
- General Purpose I/O pins
- SPI interface to Flash
- JTAG IEEE 1149.1 and IEEE 1149.6

Remote Boot
- Remote boot over Ethernet
- Remote boot over iSCSI
- PXE and UEFI

* This section describes hardware features and capabilities. Please refer to the driver and firmware release notes for feature availability.

Table 1 - Part Numbers and Descriptions

OPN	Description	Dimensions w/o Bracket
MCX4111A-XCAT	ConnectX-4 Lx EN network interface card, 10GbE single-port SFP28, PCIe3.0 x8, tall bracket	14.2cm x 6.9cm (Low Profile)
MCX4121A-XCAT	ConnectX-4 Lx EN network interface card, 10GbE dual-port SFP28, PCIe3.0 x8, tall bracket	
MCX4111A-ACAT	ConnectX-4 Lx EN network interface card, 25GbE single-port SFP28, PCIe3.0 x8, tall bracket	
MCX4111A-ACUT	ConnectX-4 Lx EN network interface card, 25GbE single-port SFP28, PCIe3.0 x8, UEFI Enabled, tall bracket	
MCX4121A-ACAT	ConnectX-4 Lx EN network interface card, 25GbE dual-port SFP28, PCIe3.0 x8, tall bracket	
MCX4121A-ACUT	ConnectX-4 Lx EN network interface card, 25GbE dual-port SFP28, PCIe3.0 x8, UEFI Enabled, tall bracket	
MCX4131A-BCAT	ConnectX-4 Lx EN network interface card, 40GbE single-port QSFP28, PCIe3.0 x8, tall bracket	
MCX4131A-GCAT	ConnectX-4 Lx EN network interface card, 50GbE single-port QSFP28, PCIe3.0 x8, tall bracket	

NOTE: All tall-bracket adapters are shipped with the tall bracket mounted and a short bracket as an accessory.

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com

53931PB
Rev 1.9